# Why Sugeno $\lambda$-Measures


Hung T. Nguyen
Department of mathematical Science
New Mexico State University
Las Cruces, NM 88003
and Faculty of Economics
Chiang Mai University, Thailand
hunguyen@nmsu.edu

Vladik Kreinovich, Joe Lorkowski, and Saiful Abu
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968
vladik@utep.edu, lorkowski@computer.org
sabu@miners.utep.edu



***Abstract*—To describe expert uncertainty, it is often useful to go beyond additive probability measures and use non-additive (fuzzy) measures. One of the most widely used classes of such measures is the class of Sugeno $\lambda$-measures. Their success is somewhat paradoxical, since from the purely mathematical viewpoint, these measures are – in some reasonable sense – equivalent to probability measures. In this paper, we explain this success by showing that while (1) mathematically, it is possible to reduce Sugeno measures to probability measures, but (2) from the computational viewpoint, using Sugeno measures is much more efficient. We also show that among all fuzzy measures which are equivalent to probability measures, Sugeno measures (and a slightly more general family of measures) are the only ones with this efficiency property.**


## I. Formulation of the Problem

**Traditional approach: probability measures.** Traditionally, uncertainty has been described by probabilities. In mathematical terms, probabilistic information about events from some set $X$ of possible events is usually described in terms of a *probability measure*, i.e., a function $p(A)$ that maps some sets $A \subseteq X$ into real numbers from the interval $[0, 1]$.

The probability $p(A)$ of a set $A$ is usually interpreted as the frequency with which events from the set $A$ occur in real life. In this interpretation, if we have two disjoint sets $A$ and $B$ with $A \cap B = \emptyset$, then the frequency $p(A \cup B)$ with which the events from $A$ or $B$ happen is equal to the sum of the frequencies $p(A)$ and $p(B)$ corresponding to each of these sets.

This property of probabilities measures is known as *additivity*: if $A \cap B = \emptyset$, then

$$p(A \cup B) = p(A) + p(B). \tag{1.1}$$

**Need to do beyond probability measures.** Since the appearance of fuzzy sets (see, e.g., [7], [10], [14]), it has become clear that to adequately describe expert knowledge, we often need to go beyond probabilities. In general, instead of probabilities, we have the expert's *degree of confidence* $g(A)$ that an event from the set $A$ will actually occur.

Clearly, something should occur, so $g(\emptyset) = 0$ and $g(X) = 1$. Also, it is reasonable to take into account that the larger the set, the more confident we are that an event from this set will occur, i.e., $A \subseteq B$ implies $g(A) \le g(B)$. Functions $g(A)$ that satisfy these properties are known as *fuzzy measures.*

**Sugeno $\lambda$-measures.** M. Sugeno, one of the pioneers of fuzzy measures, introduced a specific class of fuzzy measures which are now known as *Sugeno $\lambda$-measures* [11]. Measures from this class are close to the probability measures in the following sense: similarly to the case of probability measures, if we know $g(A)$ and $g(B)$ for two disjoint sets, we can still reconstruct the degree $g(A \cup B)$. The difference is that this reconstructed value is no longer the sum $g(A) + g(B)$, but a slightly more complex expression.

To be more precise, Sugeno $\lambda$-measures satisfy the following property: if $A \cap B = \emptyset$, then

$$g(A \cup B) = g(A) + g(B) + \lambda \cdot g(A) \cdot g(B), \tag{1.2}$$

where $\lambda > -1$ is a real-valued parameter.

When $\lambda = 0$, the formula (1.2) corresponding to the Sugeno measure transforms into the additivity formula (1.1) corresponding to the probability measure. From this viewpoint, the value $\lambda$ describes how close the given Sugeno measure is to a probability measure: the smaller $|\lambda|$, the closer these measures are.

**Sugeno $\lambda$-measures had many practical applications.** Sugeno measures are among the most widely used fuzzy measures; see, e.g., [2], [12], [13] and references therein.

*Comment.* Of course, not all expert reasoning can be described by Sugeno measures. In many practical situations, the expert's degree of confidence $g(A \cup B)$ in the event $A \cup B$ is *not* uniquely determined by the values $g(A)$ and $g(B)$. In such situations, we need to consider more general classes of fuzzy measures [2], [12], [13].

**Problem.** This practical success is somewhat paradoxical. Indeed:

- The main point of using fuzzy measures is to go beyond probability measures.
- On the other hand, Sugeno $\lambda$-measures are, in some reasonable sense, equivalent to probability measures (see [9] and Section 2).

How can we explain this?

**What we do in this paper.** In this paper, we explain the seeming paradox of Sugeno $\lambda$-measures as follows:

- Yes, from the purely mathematical viewpoint, Sugeno measures are indeed equivalent to probability measures.
- However, from the computational viewpoint, processing Sugeno measure directly is much more computationally efficient than using a reduction to a probability measure.

We also analyze which other probability-equivalent fuzzy measures have this property: it turns out that this property holds only for Sugeno measures themselves and for a slightly more general class of fuzzy measures.

The structure of this paper is straightforward: in Section 2, following the main ideas from [9], we describe in what sense Sugeno measure is mathematically equivalent to a probability measure, in Section 3, we explain why processing Sugeno measures is more computationally efficient than using a reduction to probabilities, and in Section 4, we analyze what other fuzzy measures have this property.

## II. SUGENO $\lambda$-MEASURE IS MATHEMATICALLY EQUIVALENT TO A PROBABILITY MEASURE

**What we mean by equivalence.** According to the formula (1.2), if we know the values $a = g(A)$ and $b = g(B)$ for disjoint sets $A$ and $B$, then we can compute the value $c = g(A \cup B)$ as

$$c = a + b + \lambda \cdot a \cdot b. \tag{2.1}$$

We would like to find a 1-1 function $f(x)$ for which $p(A) \stackrel{\text{def}}{=} f^{-1}(g(A))$ is a probability measure, i.e., for which, if $c$ is obtained by the relation (2.1), then for the values

$$a' = f^{-1}(a), \quad b' = f^{-1}(b), \quad \text{and} \quad c' = f^{-1}(c),$$

we should have

$$c' = a' + b'.$$

*Comment.* As we have mentioned, $A \subseteq B$ implies both $p(A) \leq p(B)$ and $g(A) \leq g(B)$. Thus, larger probability values should lead to larger degrees of confidence. It is therefore reasonable to also require that the mapping $f(x)$ that transforms the probability $p(A)$ into the corresponding degree of confidence $g(A) = f(p(A))$ be monotonic. It should be mentioned that for continuous functions $f(x)$, monotonicity automatically follows from our requirement that $f$ is a 1-1 function.

**How to show that a Sugeno $\lambda$-measure with $\lambda \neq 0$ is equivalent to a probability measure.** Let us consider the auxiliary values $A = 1+\lambda\cdot a$, $B = 1+\lambda\cdot b$, and $C = 1+\lambda\cdot c$. From the formula (2.1), we can now conclude that

$$C = 1+\lambda\cdot(a+b+\lambda\cdot a\cdot b) = 1+\lambda\cdot a+\lambda\cdot b+\lambda^2\cdot a\cdot b. \tag{2.2}$$

One can easily check that the right-hand side of this formula is equal to the product $A \cdot B$ of the expressions $A = 1 + \lambda \cdot a$ and $B = 1 + \lambda \cdot b$. Thus, we get

$$C = A \cdot B. \tag{2.3}$$

We have a product, we need a sum. Converting from the product to the sum is easy: it is known that logarithm of the product is equal to the sum of logarithms. Thus, for the values

$$a' = \ln(A) = \ln(1 + \lambda \cdot a),$$

$$b' = \ln(B) = \ln(1 + \lambda \cdot b),$$

and

$$c' = \ln(C) = \ln(1 + \lambda \cdot c),$$

we get the desired formula

$$c' = a' + b'.$$

To get this formula, we used the inverse transformation $f^{-1}$ that transforms each value $x$ into a new value

$$x' = \ln(1 + \lambda \cdot x). \tag{2.4}$$

When $\lambda > 0$, then for $x \geq 0$, we get $1 + \lambda \cdot x \geq 1$ and thus, $x' = \ln(1 + \lambda \cdot x) \geq 0$.

When $\lambda < 0$, then for $x > 0$, we have $1 + \lambda \cdot x < 1$ and thus, $x' = \ln(1+\lambda\cdot x) < 0$. However, we want to interpret the values $x'$ as probabilities, and probabilities are always non-negative. Therefore, for $\lambda < 0$, we need to change the sign and consider

$$x' = -\ln(1 + \lambda \cdot x). \tag{2.5}$$

For these new values, (2.1) still implies that $c' = a' + b'$.

From the relations (2.4) and (2.5), we can easily find the corresponding direct transformation $x = f(x')$. Indeed, for $\lambda > 0$, by exponentiating both sides of the formula (2.4), we get $1 + \lambda \cdot x = \exp(x')$, hence

$$f(x') = \frac{1}{\lambda} \cdot (\exp(x') - 1). \tag{2.6}$$

For $\lambda < 0$, by exponentiating both sides of the formula (2.5), we get $1 + \lambda \cdot x = \exp(-x')$, hence

$$f(x') = \frac{1}{\lambda} \cdot (\exp(-x') - 1), \tag{2.7}$$

i.e., equivalently,

$$f(x') = \frac{1}{|\lambda|} \cdot (1 - \exp(-x')). \tag{2.8}$$

In both cases, we can conclude that a Sugeno $\lambda$-measure is indeed equivalent to a probability measure.

*First comment: how unique is the transformation* $f(x)$? If we have two different functions $f(x)$ and $f'(x)$ with the above property, then for each triple $(a, b, c)$ that satisfies the formula (2.1), we will have $c' = a' + b'$ and $c'' = a'' + b''$, where $a'' = (f')^{-1}(a)$, $b'' = (f')-1(b)$, and $c'' = (f')^{-1}(c)$. Thus, a mapping $x' = h(x'')$, where $h(x) \stackrel{\text{def}}{=} f^{-1}(f'(x))$, has the property that $c'' = a'' + b''$ implies $h(c'') = h(a'') + h(b'')$.

It is known that the only mappings from non-negative numbers to non-negative numbers that satisfy this property are linear functions $h(k) = k \cdot x$. Thus, once we know one such function $f(x)$, all other functions $f'(x)$ satisfy the property that $f^{-1}(f'(x)) = k \cdot x$. By applying the function $f(x)$ to both sides of this equality, we conclude that

$$f'(x) = f(k \cdot x).$$

In other words, all such functions can be obtained from each other by an appropriate linear re-scaling $x \to k \cdot x$.

*Second comment: why not also require that* $f(1) = 1$. We are looking for a function $f(x)$ that transforms the probability $p(A)$ into a fuzzy measure $g(A) = f(p(A))$. For the functions (2.6) and (2.8), for $x = 0$, we have $f(0) = 0$. This equality is in good accordance with the fact that for $A = \emptyset$, we have $p(\emptyset) = g(\emptyset) = 0$ and thus, we should have $0 = g(\emptyset) = f(p(\emptyset)) = f(0)$.

Similarly, it makes sense to consider $A = X$; in this case, we have $p(X) = g(X) = 1$ and thus, we should have $1 = g(X) = f(p(X)) = f(1)$, i.e., $f(1) = 1$. Let us show that we can use the above non-uniqueness to satisfy this additional property. Indeed, once we have found the function $f(x)$, any function $f'(x) = f(k \cdot x)$ also has the desired property. We can therefore select $k$ so that we will have the desired property $f'(1) = f(k \cdot 1) = 1$. This equality means $f(k) = 1$, so we should choose $k = f^{-1}(1)$.

Specifically, for $\lambda > 0$, we have $f^{-1}(x) = \ln(1 + \lambda \cdot x)$, so $k = \ln(1 + \lambda)$, and thus,

$$f'(x') = \frac{1}{\lambda} \cdot (\exp(\ln(1 + \lambda) \cdot x') - 1) = \frac{1}{\lambda} \cdot \left((1 + \lambda)^{x'} - 1\right).$$

Similarly, for $\lambda < 0$, we have $f^{-1}(x) = -\ln(1 - |\lambda| \cdot x)$, so $k = -\ln(1 - |\lambda|)$, and thus,

$$f'(x') = \frac{1}{|\lambda|} \cdot (1 - \exp(\ln(1 - |\lambda|) \cdot x')) =$$
$$\frac{1}{|\lambda|} \cdot \left(1 - (1 - |\lambda|)^{x'}\right).$$

**So why do we need Sugeno measures?** Because of the equivalence, we can view the values of the Sugeno measure as simply re-scaling probabilities $g(A) = f(p(A))$ for the corresponding probability measure.

So why not just store the corresponding probability values $p(A)$? In other words, why not just re-scale all the values $g(A)$ into the corresponding probability values $p(A) = f^{-1}(g(A))$? At first glance, this would be a win-win arrangement, because once we do this re-scaling, we can simply use known probabilistic techniques.

## III. Processing Sugeno Measures Directly Is More Computationally Efficient than Using a Reduction to Probabilities

**What we plan to do.** Let us show that:

- while from the purely *mathematical* viewpoint, a Sugeno $\lambda$-measure is equivalent to a probability measure,
- from the *computational* viewpoint, the direct use of Sugeno measures is much more efficient.

To explain this advantage, let us clarify what we mean by direct use of Sugeno measure and what we mean by an alternative of using a reduction to a probability measure.

**The corresponding computational problem: a brief description.** We are interested in understanding the degree of possibility of different sets of events. These degrees $g(A)$ come from an expert.

Theoretically, we could ask the expert to provide us with the values $g(A)$ corresponding to all possible sets $A$, but this would require an unrealistically large number of questions.

A feasible alternative is to elicit some values $g(A)$ from the experts and then use these values to estimate the missing values $g(A)$. A possibility of such estimation follows from the definition of a Sugeno $\lambda$-measure. Namely, once we know the values $g(A)$ and $g(B)$ corresponding to a two disjoint sets $A$ and $B$, we do not need to additionally elicit, from this expert, the degree $g(A \cup B)$: this degree can be estimated based on the known values $g(A)$ and $g(B)$.

Let us explain how the desired degree $g(A \cup B)$ can be estimated.

**Estimating $g(A \cup B)$ by a directly use of Sugeno measure.** The first alternative is to simply estimate the degree $g(A \cup B)$ by applying the formula (1.2).

**Estimating $g(A \cup B)$ by a reduction to a probability measure.** An alternative idea – which is likely to be used by a practitioner accustomed to the probabilistic data processing – is to use the above-described reduction to a probability measure. Namely:

- first, we use the known reduction to find the corresponding values of the probabilities
$$p(A) = f^{-1}(g(A)) \text{ and } p(B) = f^{-1}(g(B));$$
- then, we add these probabilities to get
$$p(A \cup B) = p(A) + p(B);$$
- finally, we re-scale this resulting probability back into degree-of-confidence scale by applying the function $f(x)$ to this value $p(A \cup B)$, i.e., we compute
$$g(A \cup B) = f(p(A \cup B)).$$

**Direct use of Sugeno measure is computationally more efficient.** If we directly use Sugeno measure, then all we need to do is add and multiply. Inside a computer, both addition and multiplication are very directly hardware supported and therefore very fast.

In contrast, the use of reduction to probability measures requires that we compute the value of logarithm (to compute $f^{-1}(x)$) and exponential function (to compute $f(x)$). These computations are much slower than elementary arithmetic operations.

Thus, the direct use of Sugeno measure is definitely much more computationally efficient.

*Comment.* Of course, this result does not mean that we *should* always use Sugeno measures: for example, if we use the operation once and off-line, spending one microsecond more to compute $\exp$ and $\ln$ is not a big deal. However, in time-critical situations, with limited computational ability – e.g., if we are embedding some AI abilities into a control chip – saving computation time is important.

**How to explain the use of Sugeno measure to a probabilist.** The above argument enables us to explain the use of Sugeno measure to a person who is either skeptical about (or unfamiliar with) fuzzy measures. This explanation is as follows.

We are interested in expert estimates of probabilities of different sets of events. It is known that expert estimates of the probabilities are biased (see, e.g., [5], [8]): the expert's subjective estimates $g(A)$ of the corresponding probabilities $p(A)$ are equal to $g(A) = f(p(A))$ for an appropriate re-scaling function $f(A)$.

In this case, a natural ideas seems to be:

- to re-scale all the estimates back into the probabilities, i.e., to estimate these probabilities $p(A)$ as
$$p(A) = f^{-1}(g(A)),$$
and then
- to use the usual algorithms to process these probabilities.

In particular, if we know the expert's estimates $g(A)$ and $g(B)$ corresponding to two disjoint sets $A$ and $B$, and we want to predict the expert's estimate $g(A \cup B)$ corresponding to their union, then we:

- first, re-scale the values $g(A)$ and $g(B)$ into the unbiased probability scale, i.e., compute
$$p(A) = f^{-1}(g(A)) \text{ and } p(B) = f^{-1}(g(B));$$
- then, we compute
$$p(A \cup B) = p(A) + p(B);$$
- finally, we estimate $g(A \cup B)$ by applying the corresponding biasing function $f(x)$ to the resulting probability:
$$g(A \cup B) = f(p(A \cup B)).$$

It turns out that for some biasing functions $f(x)$, it is computationally more efficient *not* to re-scale into probabilities, but to store and process the original biased values $g(A)$. This is, in effect, the essence of applications of a Sugeno $\lambda$-measure are about.

## IV. Which Other Fuzzy Measures Have This Property: an Explanation Why Sugeno Measures Are So Practically Successful

**Reminder: what are we looking for.** We want to describe fuzzy measures which are mathematically equivalent to probability measures, but in which processing the fuzzy measure directly is more efficient than using the corresponding reduction to a probability measure.

Equivalence of a fuzzy measure $g(A)$ to a probability measure $p(A)$ means that $g(A) = f(p(A))$ for some 1-1 function $f(x)$. For the empty set $A = \emptyset$, for both measures, we should have zeros $p(\emptyset) = g(\emptyset) = 0$, so we should have $f(0) = 0$.

In general, for such fuzzy measures, if we know the values $g(A)$ and $g(B)$ corresponding to two disjoint sets $A$ and $B$, then we can compute the measure $g(A \cup B)$ as follows:

- first, we use the known values $g(A)$ and $g(B)$ and the relations
$$g(A) = f(p(A)) \text{ and } g(B) = f(p(B))$$
to reconstruct the corresponding probability values
$$p(A) = f^{-1}(g(A)) \text{ and } p(B) = f^{-1}(g(B)),$$
where $f^{-1}(x)$ denotes the inverse function;
- then, we add the resulting values $p(A)$ and $p(B)$, and get
$$p(A \cup B) = p(A) + p(B);$$
- finally, we transform this probability back into the fuzzy measure, as
$$g(A \cup B) = f(p(A \cup B)).$$

The resulting value can be described by an explicit expression
$$g(A \cup B) = f(f^{-1}(g(A)) + f^{-1}(g(B))),$$
i.e., an expression of the type
$$g(A \cup B) = F(g(A), g(B)),$$
where
$$F(a, b) \stackrel{\text{def}}{=} f(f^{-1}(a) + f^{-1}(b)) \quad (4.1)$$

For $a = b = 0$, due to $f(0) = 0$, we get $F(0, 0) = 0$.

We are looking for situations in which the direct computation of a function $F(a, b)$ is more computationally efficient than using the above three-stage scheme. Thus, we are looking for situations in which the corresponding function $F(a, b)$ can be computed fast. In the computer, the fastest elementary operations are the hardware-supported ones: addition, subtraction, multiplication, and division. So, we should be looking for the functions $F(a, b)$ that can be computed by using only these four arithmetic operations.

In mathematical terms, functions that can be computed from the unknowns and constants by using only addition, subtraction, multiplication, and division are known as *rational functions* (they can be always represented as ratios of two polynomials). In these terms, we are looking for situations in which the corresponding aggregation function is rational.

**Natural properties of the aggregation function.** From the formula (4.1), we can easily conclude that the operation $F(a, b)$ is commutative: $F(a, b) = F(b, a)$.

One can also easily check that this operation is associative. Indeed, by (4.1), we have
$$F(F(a, b), c) = f(f^{-1}(F(a, b)) + f^{-1}(c)).$$
From (4.1), we conclude that
$$f^{-1}(F(a, b)) = f^{-1}(a) + f^{-1}(b).$$

Thus,

$$F(F(a,b),c)=f(f^{-1}(a)+f^{-1}(b)+f^{-1}(c)).$$

The right-hand side does not change if we change the order of the elements $a$, $b$, and $c$. Thus, we have

$$F(a,F(b,c))=f(f^{-1}(a)+f^{-1}(b)+f^{-1}(c)),$$

i.e., $F(F(a,b),c)=F(a,F(b,c))$. So, the operation $F(a,b)$ is indeed associative.

Since we are looking for rational functions $F(a,b)$ for which $F(0,0)=0$, we are thus looking for rational commutative and associative operations $F(a,b)$ for which $F(0,0)=0$.

**There is a known classification of all rational commutative associative binary operations.** A classification of all possible rational commutative associative operations is known; it is described in [4]. Namely, the authors of [4] show that each such operation is "isomorphic" to either $x+y$ or $x+y+x\cdot y$, in the sense that there exists a fractional-linear transformation $a\to t(a)$ for which either $F(a,b)=t^{-1}(t(a)+t(b))$ or $F(a,b)=t^{-1}(t(a)+t(b)+t(a)\cdot t(b))$.

In other words, $F(a,b)=c$ means either than $t(c)=t(a)+t(b)$ or that $t(c)=t(a)+t(b)+t(a)\cdot t(b)$.

*Comment.* It should be mentioned that the paper [4] calls this relation by a fractional-linear transformation *equivalence*, not isomorphism; we changed the term since we already use the term "equivalence" in a different sense.

**Let us use this known result.** Let us use this result to classify the desired operations $F(a,b)$. First, we want an operation for which $F(0,b)=b$ for all $b$. In terms of $t$, this means that either $t(b)=t(b)+t(0)$ for all $b$, or $t(b)=t(b)+t(0)+t(0)\cdot t(b)$ for all $b$. In both cases, this implies that $t(0)=0$. Thus, $t(a)$ is a fractional-linear function for which $t(0)=0$.

A general fractional-linear function has the form

$$t(a)=\frac{p+q\cdot a}{r+s\cdot a}.$$

The fact that $t(0)=0$ implies that $p=0$, so we get

$$t(a)=\frac{q\cdot a}{r+s\cdot a}. \tag{4.2}$$

Here, we must have $r\neq 0$, because otherwise, the right-hand side of this expression is simply a constant $q/s$ and not an invertible transformation. Since $r\neq 0$, we can divide both the numerator and the denominator of this expression by $r$ and get a simplified formula

$$t(a)=\frac{A\cdot a}{1+B\cdot a}, \tag{4.3}$$

where we denoted

$$A\stackrel{\text{def}}{=}\frac{q}{r}\text{ and }B\stackrel{\text{def}}{=}\frac{s}{r}.$$

For this transformation, the inverse transformation can be obtained from the fact that

$$a'=\frac{A\cdot a}{1+B\cdot a}$$

implies

$$\frac{1}{a'}=\frac{1+B\cdot a}{A\cdot a}=\frac{1}{A\cdot a}+\frac{B}{A}.$$

Thus,

$$\frac{1}{A\cdot a}=\frac{1}{a'}-\frac{B}{A},$$

so

$$A\cdot a=\frac{1}{\dfrac{1}{a'}-\dfrac{B}{A}}$$

and

$$a=\frac{\dfrac{1}{A}}{\dfrac{1}{a'}-\dfrac{B}{A}}=\frac{a'}{A-B\cdot a'}. \tag{4.4}$$

So, for operations equivalent to $x+y$, we get

$$c'=a'+b'=t(a)+t(b)=\frac{A\cdot a}{1+B\cdot a}+\frac{A\cdot b}{1+B\cdot b}.$$

Thus,

$$c=F(a,b)=t^{-1}(c')=\frac{\dfrac{A\cdot a}{1+B\cdot a}+\dfrac{A\cdot a}{1+B\cdot b}}{A-\dfrac{A\cdot B\cdot a}{1+B\cdot a}-\dfrac{A\cdot B\cdot b}{1+B\cdot b}}.$$

Dividing both numerator and denominator by the common factor $A$, we get

$$F(a,b)=\frac{\dfrac{a}{1+B\cdot a}+\dfrac{b}{1+B\cdot b}}{1-\dfrac{B\cdot a}{1+B\cdot a}-\dfrac{B\cdot b}{1+B\cdot b}}. \tag{4.5}$$

Bringing the sums in the numerator and in the denominator to the common denominator and taking into account that this common denominator is the same for numerator and denominator of the expression (4.5), we conclude that

$$F(a,b)=\frac{a\cdot(1+B\cdot b)+b\cdot(1+B\cdot a)}{(1+B\cdot a)\cdot(1+B\cdot b)-B\cdot a-B\cdot b}=$$

$$\frac{a+b+2B\cdot a\cdot b}{1+B\cdot a+B\cdot b+B^2\cdot a\cdot b-B\cdot a-B\cdot b}.$$

Finally, by cancelling equal terms in the denominator, we get the final formula

$$F(a,b)=\frac{a+b+2B\cdot a\cdot b}{1+B^2\cdot a\cdot b}. \tag{4.6}$$

For operations equivalent to $x+y+x\cdot y$, we similarly get

$$c'=a'+b'+a'\cdot b'=t(a)+t(b)+t(a)\cdot t(b)=$$

$$\frac{A\cdot a}{1+B\cdot a}+\frac{A\cdot b}{1+B\cdot b}+\frac{A^2\cdot a\cdot b}{(1+B\cdot a)\cdot(1+B\cdot b)}.$$

If we bring these terms to a common denominator, we get

$$c'=\frac{A\cdot a\cdot(1+B\cdot b)+A\cdot b\cdot(1+B\cdot a)+A^2\cdot x\cdot y}{(1+B\cdot a)\cdot(1+B\cdot b)}=$$

$$\frac{A\cdot(a+b+(2B+A)\cdot a\cdot b)}{(1+B\cdot a)\cdot(1+B\cdot b)}.$$

Therefore,

$$F(a,b) = t^{-1}(c') = \frac{\dfrac{A \cdot (a + b + (2B + A) \cdot a \cdot b)}{(1 + B \cdot a) \cdot (1 + B \cdot b)}}{A - \dfrac{A \cdot B \cdot (a + b + (2B + A) \cdot a \cdot b)}{(1 + B \cdot a) \cdot (1 + B \cdot b)}}.$$

Dividing both the numerator and the denominator of this expression by $A$, we conclude that

$$F(a,b) = \frac{\dfrac{a + b + (2B + A) \cdot a \cdot b}{(1 + B \cdot a) \cdot (1 + B \cdot b)}}{1 - \dfrac{B \cdot (a + b + (2B + A) \cdot a \cdot b)}{(1 + B \cdot a) \cdot (1 + B \cdot b)}}.$$

By bringing the difference in the denominator to the common denominator, we get

$$F(a,b) = \frac{N}{D},$$

where

$$N \stackrel{\text{def}}{=} a + b + (2B + A) \cdot a \cdot b$$

and

$$D \stackrel{\text{def}}{=} 1 + B \cdot a + B \cdot b + B^2 \cdot a \cdot b - B \cdot a - B \cdot b - B \cdot (2B + A) \cdot a \cdot b =$$
$$1 - B \cdot (B + A) \cdot a \cdot b.$$

Thus

$$F(a,b) = \frac{a + b + (2B + A) \cdot a \cdot b}{1 - B \cdot (B + A) \cdot a \cdot b}. \quad (4.7)$$

*Comment.* Similarly to the case of Sugeno measure, we can always impose an additional requirement $f(1) = 1$, by replacing the original re-scaling $f(x)$ with a modified re-scaling $f'(x) \stackrel{\text{def}}{=} f(k \cdot x)$ with $k = f^{-1}(1)$ (for which $f(k) = 1$).

**Conclusion.** We consider the fuzzy measures $g(A)$ which are equivalent to probability measures. For such fuzzy measures, once we know the values $g(A)$ and $g(B)$ for two disjoint sets $A$ and $B$, we can compute the degree $d(A \cup B)$ as $d(A \cup B) = F(g(A), g(B))$ for the corresponding aggregation operation $F(a,b)$.

This value $g(A \cup B)$ can be computed in two different ways:

- we can reduce the problem to the probability measures, i.e., compute the corresponding probabilities, add them up, and use this sum $p(A)+p(B)$ to compute the desired value $g(A \cup B)$;
- alternatively, we can compute the value $g(A \cup B)$ directly, as $F(g(A), g(B))$.

We are looking for operations for which the direct use of fuzzy measures is computationally faster, i.e., in precise terms, for which the aggregation operation can be computed by using fast (hardware supported) elementary arithmetic operations. It turns out that the only such operations are operations (4.6) and (4.7) corresponding to different values of $A$ and $B$.

By using these operations, we thus get a class of fuzzy measures that naturally generalizes Sugeno $\lambda$-measures. Let us hope that fuzzy measures from this class will be as practically successful as Sugeno $\lambda$-measures themselves.

When do the original Sugeno $\lambda$-measures lie in this class? To understand it, let us recall that not all arithmetic operations require the same computation time. Indeed, addition is the simplest operation. Multiplication is, in effect, several additions, so multiplication take somewhat longer. Division requires several iterations, so it takes the longest time. So, any computation that does not include division is much faster. Of our formulas (4.6) and (4.7), the only cases when we do not use division are cases when $B = 0$, i.e., cases when we have $F(a,b) = a + b$ (corresponding to probability measures) and $F(a,b) = a + b + A \cdot a \cdot b$ corresponding to Sugeno $\lambda$-measures. From this viewpoint, Sugeno $\lambda$-measures are the ones for which the direct use of the fuzzy measure has the largest computational advantage over the reduction to probability measures.

*Comment.* Our result is similar to the known result that the only rational t-norms and t-conorms are Hamacher operations; see, e.g., [1], [3], [6]. The difference is in our analysis, we do not assume that the aggregation operation corresponding to the fuzzy measure is a t-conorm: for example for the Sugeno aggregation operation, $F(1,1) = 1 + 1 + \lambda = 2 + \lambda > 1$, while for t-conorm, we always have $F(1,1) = 1$.

Also, the result from [4] that we use here does not depend on the use of real numbers, it is true for any field – e.g., for subfields of the field of real numbers (such as the field of rational numbers) or for super-fields (such as fields that contain infinitesimal elements).


## Acknowledgment

This work was supported in part by the Faculty of Economics of Chiang Mai University and by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

The authors are thankful to George J. Klir and Michio Sugeno for valuable discussions, and to the anonymous referees for valuable suggestions.


## References


[1] C. Alsina, M. J. Frank, and B. Schweizer, *Associative Functions: Triangular Norms and Copulas*, World Scientific, Singapore, 2006.

[2] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer, New York, 2007.

[3] D. Butnariu and E. P. Klement, *Triangular Norm-Based Measures and Games wih Fuzzy Coalitions*, Kluwer, Dordrecht, Netherlands, 1993.

[4] J. V. Brawley, S. Gao, and D. Mills, "Associative rational functions in two variables", In: D. Jungnickel and H. Niederreiter, *Finite Fields and Applications 2001*, Proceedings of the Fifth International Conference on Finite Fields and Applications $F_q$5, Augsburg, Germany, August 2–6, 1999, Springer, Berlin, Heidelberg, 2001, pp. 43–56.

[5] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus, and Giroux, New York, 2011.

[6] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*, Kluwer, Dordrecht, 2000.

[7] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[8] J. Lorkowski and V. Kreinovich, "Likert-type fuzzy uncertainty from a traditional decision making viewpoint: how symmetry helps explain human decision making (including seemingly irrational behavior)", *Applied and Computational Mathematics*, 2014, Vol. 13, No. 3, pp. 275–298.

[9] Y. Narukawa and V. Torra, "On distorted probabilities and m-separable fuzzy measures", *International Journal of Approximate Reasoning*, 2011, Vol. 52, pp. 1325–1336.

[10] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[11] M. Sugeno, *Theory of Fuzzy Integrals and Its Applications*, Ph.D. Dissertation, Tokyo Institute of Technology, 1974.

[12] H. Tahani and J. Keller, "Information fusion in computer vision using the fuzzy integral", *IEEE Transactions on Systems, Man and Cybernetic*, 1990, Vol. 20, No. 3, pp. 733–741.

[13] Z. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 2010.

[14] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.